THE

AMERICAN NATURALIST

VOL. XLIII July, 1909 No. 511

SELECTION INDEX NUMBERS AND THEIR USE IN BREEDING [1]

DR. RAYMOND PEARL AND FRANK M. SURFACE

AGRICULTURAL EXPERIMENT STATION, UNIVERSITY OF MAINE

I. THE NATURE OF THE SELECTION PROCESS

WHATEVER one's views regarding the relative importance of different theories as to the method of inheritance for practical breeding, it must be recognized that the vast majority of all the actual breeding operations which are being carried on by plant and animal breeders at the present time are conducted by some method of selection. The belief is firmly held by most practical breeders that the bulk of the improvement which has been made in domestic plants and animals has been as the result of selection. While it is also coming to be recognized that hybridization may play an extremely important part in breeding operations either by causing increased variation or by bringing about a recombination of characters, according to Mendelian principles, yet selection must always be used as a supplement to hybridizing in practical breeding. The higher degrees of perfection demanded in judging ring and show room are only attained by the most careful and close selection. The manner in which successful selection operates, that is whether by isolation of pure or homozygote strains from a mixed population or by the gradual cumulative effect of adding together slight

[1] Papers from the Biological Laboratory of the Maine Agricultural Experiment Station, Orono, Maine. No. 11.

fluctuating variations, need not here concern us. In either case the breeder is continually meeting the same practical problem of selecting breeding stock with reference to several characters.

It is an obvious fact that a breeder practically never wishes to improve only one single characteristic of the plant or animal which he is breeding. What is usually desired is to improve several characteristics at the same time. Thus, with dairy cattle, while the main object in breeding is to increase the amount and quality of the milk other things such as constitutional vigor, breeding capacity and the like can not be lost sight of in making the selections of breeding stock. Or in maize breeding, to take an illustration from the plant side, while one may be desirous of increasing the protein content of maize, in breeding for it he must always keep in mind the conformation of the ear, size of ear, yield and a whole series of other characteristics.

While it is generally true that one wishes in practical breeding to improve more than one feature at the same time, it is an extremely difficult thing to make concurrently close selection of two or more characteristics of an organism. This difficulty is essentially a psychological one. It is the difficulty of trying to do more than one thing with the mind at the same time. The way in which this operates in breeding by the method of selection may be illustrated by an example. Suppose one is attempting to improve a strain of maize with respect to (*a*) earliness and to (*b*) conformation of the ear at the same time. When beginning his field selection of plants he makes a resolution that he will keep to a certain standard with reference to both of these characteristics and will accept nothing below those standards. Presently he comes to a plant which is by far the earliest in the field. It by a great deal surpasses all others in respect to this character, yet unfortunately the ear of this plant falls below his chosen standard for conformation of ear. What is to be done? Logically the plant ought to be rejected. But

if the breeder is deeply interested in improving the corn with respect to earliness what actually will happen will be this: he will decide to take this plant on account of its earliness and in spite of the defects of the ear. In selection work every degree of compromise of this kind is made and the larger the number of characters one attempts to deal with at the same time the more compromises are made. Where shall the line be drawn beyond which further compromise shall not be permissible?

II. The Theory of Selection Index Numbers

It would seem to be highly desirable to devise, if possible, some method for the use of the breeder who is practising selection which would get over in a greater or less degree the difficulty which has been outlined in the preceding section. What is needed is some method whereby a selection of several characters may be made at the same time in an absolutely impersonal and impartial manner without throwing out absolutely those individuals which are especially good with respect to a single one of the group of characters undergoing selection and mediocre or poor with respect to the other characters of the group. It has occurred to us that a way of reaching this desired end is found in the use of what we have called "selection index numbers," borrowing the terminology and something of the idea from the literature of political economy.

An "index number" in the sense here used means a single mathematical function which combines in itself the values of several independent or correlated variables. In such a function each of the variables may be weighted in any desired manner to meet the needs of the particular problem. Having decided the relative degree to which each variable shall be weighted, the index number as finally calculated gives an absolutely impartial and impersonal summing up of the total combined value or effect of the variables entering into it.

The theory of a breeding index number may be illus-

trated by a concrete example. Let the matter of improving maize by selective breeding be taken. Suppose that a breeder starts with a promising variety of yellow dent corn. This variety, while promising, has never been improved by breeding at all. The ears are only fair in respect to size and shape. The principal aim of the breeder of this corn, let us suppose, is to increase the "earliness" (*i. e.*, shorten the time of maturing), but at the same time he wishes to improve the other characteristics of the corn—size and shape of ear, relative proportion of corn on cob, and yield per acre. In making field selections of plants to furnish seed for testing out by the "ear-row" or other method, the breeder will endeavor to select with reference to as many as possible of the points enumerated above in addition to "earliness." Further it will be desired after the ears have been harvested and dried to take careful account in selecting the seed of at least the following points: Shape of ear, length of ear, circumference of ear, condition of tip and butt of the ear, kernel shape, germination and proportion of shelled corn to cob. Now if a plant happens to be especially early, even though it be relatively poor in respect to these ear characters, it is likely to be selected to furnish seed, in spite of these defects. But it is possible to devise a formula for a selection index number which shall give whatever weight may be previously agreed upon to each of these several variable characters of the corn which have been enumerated. Having settled upon a particular formula, the selection of seed then becomes essentially a purely mechanical matter so far as the characters included in the formula are concerned. The value of the index will be determined by the relative contributions from each of the included variables. If the breeder calculates such a selection index number, and takes no ears with an index below that standard, it will then be possible for him to select with reference to a series of characters in an unbiased and impartial manner.

Analytically considered a selection index is an expression of the general form

$$I_1 = \phi(x, y, p, q, \cdots, w) \tag{1}$$

where I_1 denotes the selection index and x, y, p, q, etc., are the variables upon which it desired to carry out selection. The practical question which has to be solved in forming a selection index in any given case as to what shall be the form of the function ϕ. The formula for an index should fulfill the following requirements:

1. It should be simple and easily calculated.
2. The value of the index should increase as the desirability of the individual as a breeder increases.
3. The index should be relatively more sensitive to small changes in important characters than to those in unimportant characters; that is, the variables should be differentially weighted.
4. The value of the index should decrease as undesirable characters become relatively marked.

It has seemed to the writers that to a first approximation the following general form of expression will be found to be well suited for a selection index:

$$I_1 = \frac{ax \pm by \pm cz \pm \cdots \pm nw}{a'p \pm b'q \pm c'r \pm \cdots \pm n't} \tag{2}$$

In this expression x, y, z, $\cdots$, w are variables which become more desirable (*i. e.*, from the breeder's standpoint) as their values *increase;* whereas p, q, r, $\cdots$, t are variables which become more desirable as their values *decrease.*[2] The quantities a, b, c, $\cdots$, n, and a', b', c' $\cdots$, n' are constants to be given arbitrary values in the proportions that the different variables are to be weighted.

It will be seen that a selection index number as described above is in a sense an adjunct or supplement to a score card. The index affords a means of condensing the entire information which the score card gives into one unit which can be then dealt with individually.

[2] Cf. actual indices for poultry and corn, pp. 392 and 397 *infra.*

Selection indices of the character above described may be separated into two general categories. These categories may for the sake of convenience be designated as (*a*) fundamental and (*b*) special. A selection index is to be regarded as fundamental when it includes in its formula those characters of an organism which every breeder will wish to maintain at a high grade or improve, whatever may be the special purpose for which he is breeding. Thus, for example, in maize breeding whether the breeder is working for high protein or high oil content or earliness, he will always desire to have well-shaped ears with good tips and of good size, a high yield of shelled corn to the acre and good vitality (*i. e.*, germinative capacity) in his seed.

These and other characteristics which we need not take space here to enumerate are in a sense fundamental or universal characteristics which every breeder wishes to improve or, at least, to keep at a high standard. Similar considerations obtain in breeding dairy cattle or poultry for performance. Besides the performance in respect to milk production or egg production every breeder of these animals desires that what may be called the "breeding capacity" of his stock shall be constantly improved or, at least, not impaired. By "breeding capacity" is meant ability to produce numerous and vigorous healthy offspring, to put the matter in the broadest way. An index number which has to do with these universally desired characters may be called the fundamental selection index for each type of organism.

On the other hand, by a special selection index is to be understood one which includes those characteristics which a breeder is specially and particularly working to improve. In the case of the corn breeder working for high protein, the protein content of the ear will constitute one of his special selection indices either by itself or in combination with some other characters. In dairy cow breeding the butter fat content of the milk will be a special selection index.

What may be called the final selection index may be formed by combining mathematically the fundamental and special indices in a single index number.

III. A Selection Index for Poultry

The writers are engaged in the experimental breeding of poultry and of maize with reference to utility characters. One of the special aims of the poultry breeding work is to learn how to fix superior egg production in a strain by breeding. In the corn breeding the chief special aim is to improve a strain of sweet corn in respect to earliness, so that it may mature seed under Maine climatic conditions. In connection with this work the idea of selection index numbers has been developed. It is proposed to discuss here the particular selection indices which we are using for poultry and for sweet corn simply as an illustration of the theory of such index numbers and of their application in actual practical breeding operations. It is not our intention to maintain in the least that the indices here given are the best which could be devised or even that they are not subject to change in our own hands with further experience. They are still in the experimental stage. Our only reason for calling attention to the particular forms of indices which we are using at this time is that we believe that the general idea of selection indices as set forth in this paper may prove to be a useful one to the breeder working in other fields. The fundamental idea of a selection index number is a most elastic one. Any one may adopt and modify these indices and weight the different variables, in any way to suit his individual needs. But there is no doubt that a general plan of this kind becomes more intelligible if concrete illustrations of its operation are given. The specific selection indices now to be discussed are presented solely as illustrations of the general theory.

In the case of poultry breeding a fundamentally desirable thing, whatever may be the special object of the

breeder, is reproductive or breeding capacity in the birds. A "200-egg" hen is of very little value *as a breeder* if she is not able to produce when mated with a good male bird a fair precentage of chickens which will live.

It is very generally stated by practical poultrymen that the point on which it is most often decided whether a given commercial venture in the poultry business shall succeed or fail is the expense involved in the hatching and rearing of the chickens. The female that will produce eggs which will yield a high proportion of chickens from the eggs set, and whose chickens live through to marketable age is an extremely desirable bird from a practical standpoint.

The fundamental selection index which we have chosen for poultry relates primarily to the breeding capacity of the female. The value of this index for a particular bird can only be determined after her breeding capacity has been tested. This poultry selection index comes into application in deciding which of the pullets that have been tested as breeders in their pullet year shall be kept over to be used as breeders in their second year of life.

The formula which has been provisionally adopted in our work as a fundamental poultry selection index is as follows:

$$I_1 = \frac{5(a+b)}{c+d+1}$$

The following scheme shows the meaning of the letters in the formula:

I_1 = general or fundamental poultry selection index for an individual bird.

a = percentage of this bird's eggs which hatched.

b = percentage of eggs actually laid by this bird to the total number it was possible for her to lay between February 1 and June 1 (*i. e.*, the breeding season) of the year for which the index is calculated.

c = percentage of this bird's eggs which were infertile.

d = percentage of chicks hatched from this bird's eggs

which died within three weeks from the date of hatching.

A brief discussion will make clear the reasons why these particular variables are chosen for the index and are arranged in the formula in the manner that they are. A bird's value as a breeder increases as the percentage of her fertile eggs which hatch increases. Therefore a should go into the numerator of the index fraction. Similarly a bird's value increases as a breeder in proportion to her egg production in the breeding season. A bird which produces few eggs during the breeding season (whatever she may have done before) *ipso facto* can not produce many chickens. Instead of using the actual egg production in the index the relative or percentage production is used, for reasons which have been discussed in a previous publication by the present writers.[3]

Now in distinction to the factors so far discussed it is clear that a hen's value as a breeder decreases as the number of infertile eggs which she produces in a given time increases. To put this factor into the breeding index is, of course, equivalent to asserting that the hen plays at least an equal part with the cock in determining fertility. This is undoubtedly the case, though it is not the place here to present detailed evidence on the point. Since relatively poor fertility of the eggs is an undesirable characteristic c is put in the denominator of the index fraction. The case is the same in regard to d. If the chicks produced by a particular bird are weak and die early when given the same treatment as that under which other chicks thrive, it is an indication that that particular bird is not desirable as a breeder.

In order to give a concrete idea of the values which this poultry selection index may take in actual cases Table I has been prepared. This table gives the value of I_1 for a number of pullets which were tested as breeders in the spring of 1908, as well as the values of the component factors in each case. The cases given in the table

[3] Cf. Me. Agric. Expt. Station, Bulletin No. 165, pp. 46–48.

are not a random sample of the breeding flock but were definitely selected to illustrate the behavior of the index.

TABLE I

SHOWING REPRESENTATIVE VALUES OF THE SELECTION INDEX IN POULTRY AND ITS COMPONENT PARTS

DATA FOR BARRED PLYMOUTH ROCK PULLETS[4]

Band No. of Bird.	A Per Cent. of Fertile Eggs Hatched.	B Percentage Breeding Egg Production.	C Percentage of Eggs Infertile.	D Per Cent. of Chicks Dying within 3 Weeks.	Selection Index I_1.	Band No. of Bird.	A Per Cent. of Fertile Eggs Hatched.	B Percentage Breeding Egg Production.	C Percentage of Eggs Infertile.	D Per Cent. of Chicks Dying Within 3 Weeks	Selection ndex I_1.
10	21	18.33	36	33.33	2.8	29	28	28.33	27	0	10.0
160	9	15.83	42	0	2.9	23	39	35.83	28	8.33	10.0
402	14	33.33	30	50.00	2.9	428	46	40.08	20	22.22	10.0
352	12	41.67	14	60.00	3.6	122	49	34.17	15	23.53	10.5
358	50	31.67	32	69.23	4.0	375	41	37.50	36	0	10.6
438	35	35.00	45	37 50	4.2	712	42	50.00	20	20.00	11.2
441	38	20.00	33	33.33	4.3	408	48	45.86	16	22.73	11.8
21	25	35.00	14	44.44	5.0	38	61	37.50	9	28.00	13.0
393	12	47.50	9	50.00	5.0	731	27	35.83	23	0	13.1
705	38	26.67	34	25.00	5.4	395	29	37.50	24	0	13.3
717	24	21.67	19	20.00	5.7	443	56	54.17	6	35.29	13.3
39	23	29.17	26	16.67	6.0	409	37	46.67	4	25.00	13.9
377	32	37.50	16	33.33	6.9	771	43	60.00	13	22.22	14.2
746	59	28.33	15	47.06	6.9	19	68	24.17	24	6.66	14.5
87	36	39.17	17	35.71	7.0	152	26	39.17	11	9.09	15.5
359	61	32.50	15	50.00	7.1	366	52	25.83	13	7.14	18.4
442	41	68.33	38	28.57	8.2	768	74	45.00	20	9.38	19.6
400	44	47.50	16	38.10	8.2	434	72	58.33	17	14.28	20.2
27	18	40.00	8.3	25.00	8.5	750	57	52.50	16	10.00	20.3
757	23	51.67	10	30.77	8.9	770	71	50.83	9.8	12.82	25.8
725	33	29.17	6	27 27	9.1	752	48	59.17	6	12.50	27.5
112	17	46.67	18	25.00	9.3	397	38	40.00	6	5.89	30.3
753	61	47.50	46	10.53	9.4	168	88	35.83	4.7	13.89	31.6
407	41	41.67	18	23 53	9.7	749	57	46.50	4	6.45	45.2

From this table the following points are to be noted:

1. The selection index ranges in value between 2.8 and 45.2 in the cases chosen for illustration. These figures do not represent the extreme values which may be obtained for this index. They suffice, however, to bring out the point of practical importance that the index fluctuates through a wide numerical range as the value of the birds as breeders changes.

[4] It should be said that all the records given in this table were made under uniform conditions as to housing, feed, age of birds, etc. The hatching was all done in incubators.

2. A study of the detailed figures of the table makes it apparent that the index is a perfectly impartial and accurate measure of a bird's breeding performance. Thus to take an example, bird no. 402 has a low selection index (2.9), while bird no. 168 has an index more than ten times larger (31.6). Do these figures represent fairly the difference between these two birds in respect to breeding performance? Let us examine the detailed figures. But 14 per cent. of the fertile eggs from bird no. 402 hatched as against 88 per cent. of the fertile eggs of no. 168. In regard to number of eggs laid during the breeding season the two birds are about equal, with what advantage there is in favor of no. 168. But on turning to the question of fertility of eggs we see that 30 per cent. of the eggs of no. 402 were infertile as against only 4.7 per cent. in the case of no. 168. The same condition obtains in regard to the vitality of the chicks. Fifty per cent. of the chicks from no. 402 died before attaining the age of three weeks, whereas but 13.89 per cent. of no. 168's chicks died within this time. There is no doubt of the great superiority of 168 over 402 in breeding performance. The index gives an exact measure of its degree or amount.

3. The detailed figures bring out clearly the further point that the value taken by the index is not unduly influenced by any one factor. Low values for one variable may be offset by high ones in another. In general, the variables which form the numerator of the index fraction are seen to increase as the index increases. The reverse holds in general for the variables in the denominator of the index. The facts regarding the two variables, a (percentage of fertile eggs hatched) and c (percentage of infertile eggs) are shown graphically in Fig. 1.

The diagram shows that the percentage of fertile eggs hatched in general increases as the selection index increases in value, but with much greater fluctuations from bird to bird. On the other hand the percentage of infertile eggs in general decreases as the value of the index increases, but again with much greater fluctuations. The

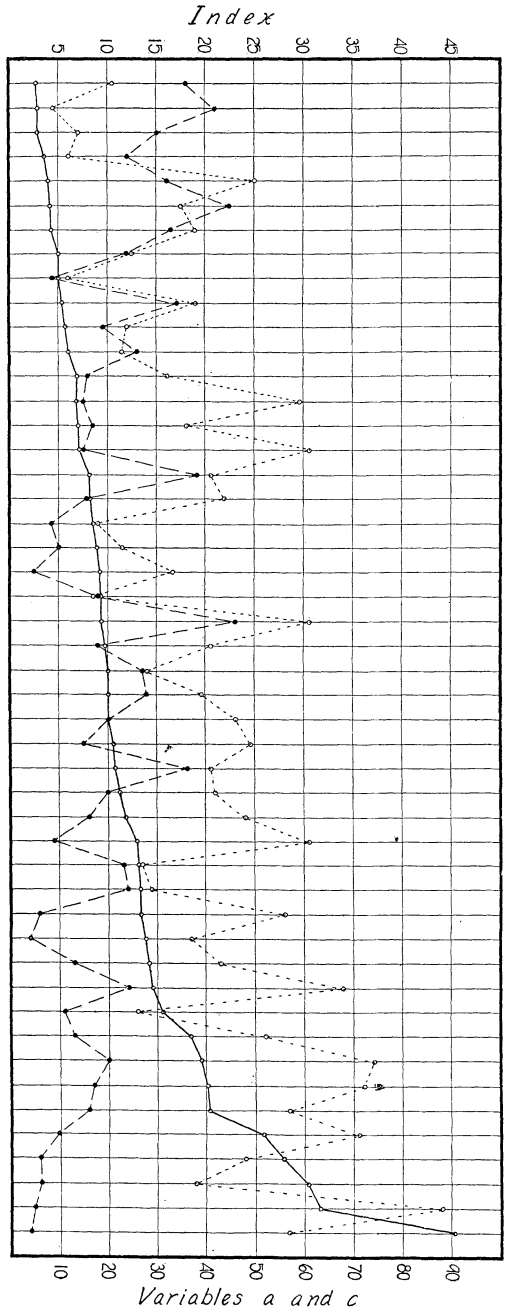

FIG. 1. Diagram showing the relation to the value of the selection index of the order given in Table I; O———O selection index; o.....o percentage of two of the variables entering into it. The abscissæ denote the birds arranged in fertile eggs hatched; ●---● percentage of infertile eggs.

smoothing effect of combining the four variables into one index number appears most clearly from the diagram.

4. In Table I, one half of the birds have a selection index equal to or greater than 10. This is about the value which it would seem desirable to take as a division point in selecting breeders. A bird with an index below 10 can not be regarded as a good breeder.

Summarizing, we believe the poultry selection index described, or some modification of it to suit particular needs, to be a useful aid in practical breeding operations with poultry. It measures in an exact and impartial manner the performance of a bird as a breeder in any given season. On the basis of the knowledge so gained the breeder can select birds which are to be retained for further breeding. It substitutes an exact and impartial measure, in the place of a rough, general impression of the relative effect of several variables.

IV. A Selection Index for Sweet Corn

The idea of selection index numbers is being applied in connection with investigations in breeding sweet corn (*Zea Mays saccharata*). In this work the selection index is based on the characters of the ear. A brief discussion of this index is introduced here for two reasons, viz., (*a*) to illustrate the adaptability of the selection index idea to widely different classes of breeding problems, and (*b*) to show how such index numbers supplement the score card in breeding operations. In planning this corn selection index the underlying idea was that set forth by way of example earlier in the paper (p. 386, *supra*).

The actual index number used for sweet corn has the following formula:

$$I_1 = \frac{A + 3B + 2C}{D + E + F^2}$$

The meaning of the letters in the equation are shown in the following scheme.

A = Length of ear multiplied by the circumference of the ear ($L \times$ Cir.). This gives a measure of the absolute size of the ear.

$$B = \frac{\text{100 times the weight of grains shelled from butt and tip of ear.}}{\text{Total weight of shelled corn from whole ear}}$$

B is the percentage which the grain on butt and tip is of total grain on the ear. This gives a measure of the extent to which ends of the ear are covered with grain.

$$C = \frac{\text{100 times the circumference of the cob at the middle of the ear.}}{\text{Circumference of ear at middle.}}$$

This measures in effect the average percentage depth of kernel (by difference).

$$D = \frac{\text{100 times the weight of the cob.}}{\text{Total weight of shelled corn from the whole ear.}}$$

E = (circumference of ear at butt) − (circumference of ear at tip). A measure of the shape of the ear.

F = 100 − the observed percentage germination of grain from this ear.

All measurements and determinations are made on the thoroughly dried ear and are recorded in centimeters and grams.

As an illustration of the way in which this index is calculated a single example may be worked out in detail.

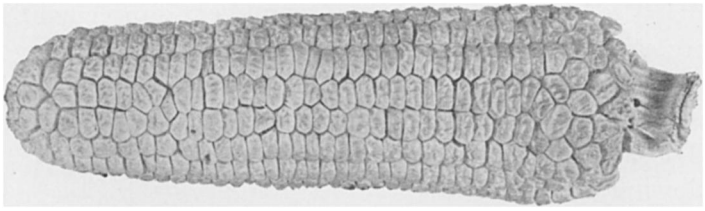

FIG. 2. Photograph of sweet corn ear No. 596 for which the selection index number is calculated in the text.

Let us determine the value of I_1 for the ear of sweet corn shown in Fig. 2, and recorded in our notes as ear No. 596.

For this ear the following measurements were recorded:

Length	14.8	cm.
Circumference	12.9	"
Weight	118.0	gr.
Weight of shelled corn	97.0	"
Weight of corn shelled from butt and tip	22.8	"
Weight of cob	21.0	"
Cob circumference	7.9	cm.
Butt circumference	9.3	"
Tip circumference	6.5	"
Rows	16	
Number of kernels in average row	33	
Germination (100 kernels tested)	100	per cent.

Using these data to calculate the selection index it appears that:

$$A = 14.8 \times 12.9 = 190.9$$

$$B = \frac{100 \times 22.8}{97.0} = 23.5$$

$$3B = 70.5$$

$$C = \frac{100 \times 7.9}{12.9} = 61.2$$

$$2C = 122.4$$

$$D = \frac{100 \times 21.0}{97.0} = 21.6$$

$$E = 9.3 - 6.5 = 2.8$$

$$F = 100 - 100 = 0$$

$$F^2 = 0$$

Whence we have:

$$I_1 = \frac{190.0 + 70.5 + 122.4}{21.6 + 2.8 + 0} = \frac{333.8}{24.4} = 15.7$$

The value of the index in this case is much above the average for the particular variety of sweet corn of which this ear is a representative. This is, of course, as it should be, since the ear is a particularly fine one for sweet corn. Its aristocratic lineage is apparent from the photograph. Some ears of this variety give a value for the index of as low as 1.0 or even lower. It is obvious that the index number for an ear of one variety is not directly comparable with that for an ear from another

variety. Corn index numbers, as formulated above, are only directly comparable within a single variety. They may of course just as well be used in judging a dent corn, for example, as for the sweet corn here discussed. The average ear of dent corn will obviously give a different value for the index from that given by the average ear of sweet corn.

Any one familiar with the score cards used in judging corn in the various agricultural shows and fairs in this country will recognize that the corn selection index formula here given combines in one expression quantitative determinations of a majority of the characters which appear on the score card. The index in a way takes the place of the judge. It impartially "cuts for each defect" according to a previously agreed upon system of weights. By the combined use of the index number, scales and measuring tape, unconscious or conscious partiality and bias is inevitably and absolutely taken out of the judging. The selection index idea seems capable, when properly developed to meet the needs of particular cases, of supplying in some measure that thing which has been so long desired in all kinds of stock judging, whether of plants or animals, viz., an *absolute* base or datum plane.

V. SUMMARY

The purpose of this paper is to call the attention of those interested in breeding operations to the usefulness of what we have called "selection index numbers" in such work. The idea of such index numbers is to combine in a single numerical expression the values of a series of variable characters with regard to all of which the breeder wishes to practise selection at the same time. The analytical expression of this idea is discussed and its adaptability and usefulness are illustrated by examples drawn from poultry and maize breeding. It is shown that selection index numbers form a valuable adjunct to the score card in stock judging.